DOCUMENT RESUME

ED 441 286 EA 030 414

AUTHOR Mintrop, Heinrich; MacLellan, Ann Merck; Quintero, Margaret Fee; Kelemen, Matthew

TITLE The Design of School Improvement in Schools on Probation: A Comparative Content Analysis of School Improvement Plans in Three Accountability Systems.

PUB DATE 1999-10-00

NOTE 68p.; Paper presented at the Annual Meeting of the University Council for Educational Administration (Minneapolis, MN, October 29-31, 1999).

PUB TYPE Reports - Evaluative (142) -- Speeches/Meeting Papers (150)

EDRS PRICE MF01/PC03 Plus Postage.

DESCRIPTORS *Accountability; Comparative Analysis; *Competency Based Education; *Educational Change; *Educational Improvement; Educational Technology; Elementary Secondary Education; Instructional Improvement; Management by Objectives; Motivation Techniques; *Performance Based Assessment; Power Structure; *Probationary Period; Public Schools; Systems Analysis

IDENTIFIERS Indicators; Kentucky; Maryland; San Francisco Unified School District CA

ABSTRACT

Accountability systems target effective school management among other key aspects of the school-improvement process that impact teacher performance. In treating this aspect of accountability, it asks, "What do school improvement plans reveal about schools' responses to accountability and probation?" The school-improvement plan (SIP)--usually required to codify a school's envisioned improvement design-- is a common feature among accountability systems in Maryland, Kentucky, and the city of San Francisco. This comparative analysis links specific patterns of schools' plans to specific design features of the accountability system. All three systems bring performance accountability to schools via external imposition but differ in the way they use different performance indicators, selection criteria for probation, and capacity building measures. The report also explores whether differences in the accountability systems (the presence or absence of specific design features) can be associated with differential patterns in school-improvement plans. (Contains 52 references.) (DFR)

Reproductions supplied by EDRS are the best that can be made from the original document.

ED 441 286

The Design of School Improvement in Schools on Probation

A Comparative Content Analysis of School Improvement Plans in Three Accountability Systems

Heinrich Mintrop
Ann Merck MacLellan
Margaret Fee Quintero
Matthew Kelemen

University of Maryland
University of Kentucky
Stanford University

Paper presented at the Annual UCEA Conference,
Minneapolis 1999

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
H. Mintrop
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

EA 030 414

BEST COPY AVAILABLE

In recent years, school district administrations and state governments all over the U.S. have begun to install school accountability systems that are aimed to improve the productivity of school organizations. Most accountability systems consist of standards for school and student performance, indicators to measure those standards, public exposure and often scrutiny of schools' performance data, and rewards and sanctions for schools that either excel or fail according to the systems' performance criteria. Schools that persistently perform below the expectations of the accountability agency are a special case in accountability systems. They are often subjected to a period of probation during which they are challenged to reverse their decline. During this period, in many cases, control by central administrative agencies and the provision of capacity building measures intensify. A school that successfully masters its probation will exit the stage of "high-stakes," during which the threat of sanctions is presumably imminent; a school that fails to improve despite its publicly exposed shortcomings, despite managerial procedures mandated by the accountability agency, and despite the provision of capacity building will presumably experience the full force of sanctions.

In some jurisdictions, these sanctions are associated with the term "reconstitution." Although what is actually meant by reconstitution differs somewhat by jurisdiction, reconstitution evokes the threat on the part of the accountability agency to seize control of the school's operations and to evaluate the staff's performance, and perhaps tenure, in the building or the district. In many instances, reconstitution entails "zero-based staffing," i.e. faculties are subject to losing their building seniority and are forced to reapply for their old positions. Whole faculties, rather than individual teachers, are held accountable for increasing the school's organizational performance.

Most accountability designs mandate schools, upon being identified as on-probation, to

undergo a formal process of school improvement, often stipulating the compilation of school improvement plans (SIP), formation of responsible bodies of coordinated action such as school improvement teams (SIT), and monitoring of the plan's implementation. Reconstitution designs differ in the degree to which they offer additional resources for capacity building. The assessment system presumably provides the school clear goals centering on student learning. The combination of mandates and incentives aims at motivating educators in schools on probation to increase effort, at channeling schools' activities into a rational, i.e. goal-oriented and coherent process of organizational learning and coordination, and finally at restructuring and intensifying curriculum and instruction through teacher learning and instructional change. Thus, accountability systems target three key aspects of the school improvement process that presumably impact teacher performance: educators' motivation, effective school management, and instructional technology and capacity to increase student performance as measured by the assessment instruments. In this paper our focus is narrow: we will present data on the management aspect of accountability by asking the question what school improvement plans can tell us about schools' responses to accountability and probation.

School accountability designs differ across jurisdictions with regard to the indicators for identifying and selecting failing schools, the nature of the threatened sanctions, and the extent of external monitoring and technical assistance offered to, or imposed on, schools. Our paper looks at three such designs, the states of Maryland and Kentucky, and the city of San Francisco. Key differences among them will be explained below. A common feature among all three accountability designs is the school improvement plan. Most jurisdictions that have experimented with probation require schools to prepare school improvement plans (SIP) that codify a school's envisioned improvement design. These SIP's, most often subjected to official

review and approval, are a centerpiece in accountability systems and schools' change efforts. Whether they are hastily thrown together or carefully crafted, these school improvement plans are vivid testimony to the way schools (and districts) think about the task of school improvement under unusual conditions of probation.

As public and official documents, SIPs represent the "espoused" (Schein, 1991) views of teachers and administrators on the task of improving their schools. Naturally, espoused views are not necessarily identical to educators' common-sense or more deeply held beliefs about how their school should be changed, nor are espoused designs necessarily implemented designs. The documents themselves will not unveil layers of lived culture and practice. Only case study data on school improvement plans in the context of a given school can reach these layers. But as negotiated links between official expectations of the accountability agency and practitioners' perception of school reality, these documents mark the way schools-on-probation attempt to systematize the design of improvement for their school. By identifying design patterns, documentary analysis contributes one piece of the answer to a larger inquiry into the effect of accountability systems on school improvement, here specifically on the improvement of schools identified as probationary.

Such inquiry is serviced even further if we can link specific patterns of schools' plans to specific design features of the accountability system. Our comparative analysis across the three jurisdictions attempts to do just that. All three jurisdictions are similar in that they bring performance accountability to schools via external imposition. But they differ in the way they do so. They use different performance indicators, selection criteria for probation, and capacity building measures; and they differ in the way the school improvement plan is embedded in the bureaucratic authority structure of the school system. So we not only want to find common

patterns in school improvement plans across the three jurisdictions that might be indicated by schools' common position of probation, but we also explore if differences in the accountability systems, i.e. the presence or absence of specific design features, can be associated with differential patterns in school improvement plans.

Planning in Schools

Planning is a key ingredient of the type of school improvement that takes its cues from effective schools research. Lezotte and Bancroft (1985) summarize the principles of effective schools research in the following way. School improvement begins with educators' *motives* for creating a high quality education that is equitable and serves all children. Improvement efforts are *research-based*, i.e. program designers utilize research knowledge on current best practice. Improvement is *data driven*, i.e. needs are derived from measurable and observable evidence. Data are disaggregated for various student groups and skills. School improvement concentrates on teaching and learning. Success is measured in terms of student outcomes. Student learning and behavior goals are agreed upon by the school, and progress towards goals is evaluated through frequent monitoring of student learning. The school engages in a process of organizational planning and development, informed by the belief that improvement is possible without a large infusion of resources or personnel. According to Lezotte and Bancroft (1985), major steps of the improvement process are: forming a school improvement team, setting goals, assessing needs, identifying effective practices, disseminating the plan among faculty, implementing the plan, and evaluating outcomes. School improvement plans, then, are seen as key devices to facilitate a school's refection on change and management of change.

This basic rational model of school planning consisting of needs assessment, goal formation, action plan, implementation, monitoring of activities and evaluation of outcomes is the foundation for a great variety of school improvement programs (Morrison, 1998), such as management by objectives (Bell, 1974) or organizational development of schools (Neale, 1981). Models have been suggested that carry rational organizational planning to each individual employee. In "performance management" (Mitchell, 1995), standards, indicators, goals, and objectives are formulated for the school organization and then translated into individual performance expectations. The key in performance management is the alignment of individual growth plans and school improvement plans. When the two are aligned, each step of the organizational plan becomes the responsibility of specified individuals at the school.

A similar improvement model (Tucker & Codding, 1998) has been developed in the context of the recent reform wave towards standard-based education. It concurs with the above described models in its focus on performance standards, data-based diagnostics, measurable indicators, and planning as main leverage points for school improvement. But it expands the narrow management focus by including a phase of initial cultural reorientation during which schools examine their "core beliefs," and a phase of benchmarking best practices during which educators fan out and experience teaching models that work in other settings. Tucker and Codding list six steps for their standard-based improvement model: set standards, formulate vision and mission of the school, analyze the causes of the problem, set performance targets, select benchmarked strategies, draw up action plans that specify who is responsible for what activity using which resources, implement the plan, and monitor implementation.

The disappointing results of educational policy implementation in schools in the sixties and seventies elicited a critique of rational models of school improvement with their reliance on

planning (Firestone & Corbett, 1988). Critics of the rational model see educational systems as loosely coupled (Weick, 1978) rather than tightly connected in their goals, structures, activities, and outcomes; decision making in educational organizations is anarchic, rather than based on clear means-ends relationships (Olsen, 1976), resulting in incrementalist approaches to change (Lindblom, 1959; Conley, 1993); schools have ambiguous goals and an ill-defined technology that make connections between cause and effect, instructional strategies and outcomes tenuous; rather than being masters of their own fate, schools are fundamentally shaped by their institutional environment (Meyer & Rowan, 1978); and these school environments are in a constant state of turbulence (Patterson, Purkey & Parker, 1986; Wallace, 1994); due to this institutional nexus, educators' actions in schools are strongly guided by norms and values that community, wider society, and the profession have traditionally deemed appropriate and legitimate for schools (Deal & Peterson, 1999). The organization, on the other hand, tries to ritualistically comply with legitimate models that are foisted upon it by other organizations in its environment (DiMaggio & Powell, 1983) while its core remains unresponsive to external policies (Elmore, 1996).

In the view of some, this state of affairs is deplorable because it renders schools ineffective. When goals are unclear and unfocussed and goals and activities are unconnected to identified needs, when causes for under-performance are sought in the school's external environment rather than attributed internally, when suggested activities are incremental and incoherent add-ons to the school's existing core instructional program, and when finally changes are accompanied by demands for new resources, the prospect for effective improvement diminishes. A group of economists (Hanushek, 1994) theorize that past school reform attempts have not improved student performance and have encouraged waste of human and financial

resources because schools and educators lack clear performance incentives. They concede, however, that "nobody knows what incentives work effectively in the varied settings of the nation's schools" (Hanushek, 1994:53). But some design principles for a good incentive system are suggested. A good incentive system is tightly linked to student performance. It specifies goals and leaves it up to educators to decide how to achieve them so that schools can pursue solutions that best fit their unique needs. Since the link between resources and inputs, on one hand, and student outputs, on the other hand, is weak and not clearly understood, a good incentive system balances "flexibility in the means of education" with "crystalline clarity regarding the desired ends" (Hanushek, 1994:88).

Rational models of organizational improvement have gained political impetus in recent systemic reform efforts. A press for school improvement occurs in a systemic model through the alignment of system goals with school organizational goals and through coherence of performance goals, rewards, and sanctions that authoritatively impinge on school practitioners (Smith & O'Day, 1991; O'Day & Smith, 1993). It is hoped that clear standards coupled with accountability will tighten the connection between policies and teacher behavior. For under-performing schools that fail to undertake self-correcting actions in response to performance information, sanctions may be imposed. "*High stakes* in theory will increase motivation and performance" (O'Day, 1993:286). In the logic of systemic reform, school improvement plans make sense as a management tool that rationally aligns a system's performance demands with practitioners' actions and change efforts. With an authoritative performance assessment system in place, schools are provided the external impetus to focus on student achievement and to adjust their own expectations of students to the high expectations of the system for the school's performance. Consensus on goals and standards, a postulate of early effective schools research, is

more easily attained when school improvement is embedded into an external accountability system. While systemic reform efforts, theoretically, leave it in the autonomy of the school to choose specific strategies for improvement, those activities are expected to be aligned to the goals of the accountability agency. Thus goal focus and alignment are key features of a successful accountability system.

Others accept the lack of rationality in schools as a fact or consider it a sign of a genuine quality of education that distinguishes this sphere from business and administrative rationality. For Wise (1977), the insistence of external forces on changing schools through rational means brings forth "hyperrationalization," a state in which the less formalized, goal-diffuse, and personalized world of educators is subjected to excessive proceduralization and legalization (e.g., reporting and monitoring requirements) to induce compliance. Likewise Fullan points to the same "rational fallacy" (Fullan, 1991:108-110).

The wisdom of an accountability system that exposes schools to goals in form of external demands is doubted by McDonald (1996) summarizing his experience with change processes in schools affiliated with the Coalition of Essential Schools. School change, in his view, is not "leveraged" whereby specific levers (e.g., standards) drive the improvement of all essential areas of the organization. The alternative he describes is a view of change as relational whereby faculties become involved in an on-going reflective conversation on the school's purposes, students, student work, and standards. Standards involve the whole school as a community. They are formulated in a dialogue with students and parents about the necessary steps to achieve these standards. The core source of this conversation, Meier's (1995) touted account of school change suggests, is the moral empathy of adults towards the children they have chosen to educate. Real standards, says McDonald, "have to be constructed on-site." They can be externally formulated,

but have to be substantiated "by the light of the actual performance of students" (McDonald, 1996:151), i.e. they have to be internalized. External accountability carried out by way of rigorous performance testing may actually have a detrimental effect on those schools that have begun to formulate internally generated authentic standards and to build community around them, as in Newmann's sample of restructuring schools (Newmann, King & Rigdon, 1997).

A great number of studies have shown the importance of internal development if schools are to tackle ingrained cultural regularities (Sarason, 1990), become places of innovation (Little, 1982; McLaughlin & Talbert, 1993) and engage in instructional reform that stresses authenticity, in a dialog of teaching and learning in which teachers express intellectual curiosity and personal caring and students disclose their genuine interests. Developing schools in this direction is an organic process that involves adults struggling with difficult issues and dilemmas in an atmosphere of trust and mutual commitment (Nias et al., 1992; Barth, 1990). This process includes the micro-politics of change at the school site (Muncey & McQuillan, 1996; Ball, 1990) and the culture of the school which is influenced by the artistry of the symbolic leader who validates norms and values and who reinforces rituals and ceremonies. A positive school culture not only fosters productivity, but also builds commitment, boosts energy and vitality of the staff, and focuses daily action on what is deemed important (Deal & Peterson, 1999:8).

How do schools and school systems plan under these conditions of external demands for accountability and internal vicissitudes of school development? In answering this question, we will look a three different sets of studies: a study on U.S. school districts' engagement in strategic planning (Conley, 1992;1993), an evaluation of a voluntary planning effort in a New York City based school improvement project (Clark & McCarthy, 1983; Canner, 1985), and studies on planning efforts under conditions of accountability coming out of Britain (Hargreaves

& Hopkins, 1994; Broadhead et al., 1996). Though this empirical base does not allow for generalizing statements about schools' planning, the studies illustrate a specific pattern of planning that holds across various contexts. Planning does not come natural to schools (Broadhead et al., 1996; Hutchinson, 1993). Schools left to their own devices would chart their course on a day to day basis. As a result, planning efforts seem to be introduced from the outside, usually from the top of the system (Conley, 1992; Clark & McCarthy, 1983). Entering into schools from the top, it is not surprising that the principal and her functionaries are a key presence and force in schools' planning efforts, (Broadhead et al., 1996; Constable, 1994; Canner, 1985) though principals are rarely involved in planning alone (Broadhead et al., 1996:283). Broad teacher participation in planning is not common (Biott, Easen & Atkins, 1994). One study found that a large percentage of classroom teachers did not hold a copy of their school's plan (Broadhead et al., 1996:284). Though feeling committed themselves, often planning committees have doubts about the commitment of non-committee members to the process (Canner, 1985; Biott, Easen & Atkins, 1994). Educators tend to comply with planning mandates, but resistance among them against control through rational management systems can be great (Rowan, 1986).

The produced plans have been found to be unrealistically comprehensive and full of minutiae rather than being focused and strategic (Broadhead et al., 1996; Levine & Leibert, 1987). In schools, planning efforts tend to gravitate towards documentation of grand visions and routine tasks (Conley, 1993) or tend to loose their strategic character by becoming occasions for conversations about day-to-day operations (Clark & McCarthy, 1983; Wallace 1994). Conley (1993) undertook a detailed content analysis of district plans written in the 1980's under the auspices of a wave of strategic planning. He found that philosophy and mission of the plans were

an eclectic assortment of tenets that reflected the "conventional wisdom" of the time, that goals and objectives tended to be vague or marginally realistic, that strategies were numerous and activities incrementalist, that the instructional core was not the primary focus of planning and that the action plans "did not suggest that districts were pursuing highly innovative or original approaches to curricular and instructional improvement." (Conley, 1993:23). On the other hand, he found that educators believed that planning served as an occasion to increase constituents' involvement in school improvement, i.e. the plan served either as a rallying point for communication or as public posturing. The actual function of the document as a guideline for future school activities was diminished.

According to Stone & Brush's (1996) theory on planning within organizations in ambiguous contexts, this outcome is not surprising. Stone & Brush begin with the observation that such organizations rarely plan unless forced to do so. According to this model, organizations in ambiguous contexts that have to deal with diffuse goals, the influence of multiple constituencies, and a lack of direct control over resource flows find it difficult to engage in formal planning. Relying on extra-economic psychic and social rewards to maintain the commitment of varied internal and external constituencies to the organization necessitates trade-offs among competing goals; by contrast, goal formalization may increase conflict by bringing trade-offs out in the open. The instability of resource flows, and of the organization as a whole, makes even medium-range plans obsolete quite easily. Improvisation, necessary under these conditions, is actually hampered by formal plans.

At the same time, non-profits rely on the support of larger organizations or systems (businesses, government) for support in which models of rational planning are considered a badge of good business practice. Organizations in ambiguous contexts are therefore caught in a

bind between commitment and legitimacy. They compile a plan to gain external validation and legitimacy while at the same time trying to leave room for internal maneuvering. Thus, plans have an external legitimacy function, rather than giving the organization internal direction.

Although this model was derived from studies of non-profit organizations, rather than public bureaucracies of which schools are a part, the model is applicable to schools, to a certain degree, if one considers schools under conditions of accountability and expanded management autonomy as quasi-non-profits that have to exploit their environment for attracting both resources and customers. Thus, it could be expected according to this model that planning in schools remains external to the inner workings of the organization. Accordingly, when schools are left to their own devices, self-evaluation often results in improvement plans that are superficial and defensive (Cuttance, 1994). The problem of balancing external legitimacy and internal commitment becomes sharpened when schools fall under an accountability regime that formulates goals for them externally and that treats plans not only as a signal of compliance with these goals, but also as a lever to reach the core of schools' operations in order to align core activities with the performance goals of the accountability agency. While, from the vantage point of schools, planning may take the form of a dance around issues of legitimacy and commitment, the accountability agency presumably wants to see schools translate external demands into obligations for true accountability and charge internal commitment with the task of true development, i.e. it expects schools to internalize the rationality of the accountability system and carry it to its core operations.

The problem of internalization has been recognized by many who study school planning. Some paint a picture of intense self-examination, most dramatically formulated by researchers who studied schools that were voluntary participants in school planning: "As one SIP (= School

Improvement Project) principal expressed it, joining SIP is like joining Alcoholics Anonymous. first the school must realize it needs to improve, then it must willingly plunge in and help itself." (Clark & McCarthy, 1983:20). In the absence of internalization, schools unproductively spin their wheels. In Levine and Leibert's account of planning, teachers and administrators are bogged down in an overload of activities, reporting, and paperwork. Planning "produces little more than scurrying about to provide 'evidence' that the school staff is in compliance with the planning guidelines." (1987:399; see also Bardach, 1986).

The problem of internalizing external demands for accountability and rational planning has often been addressed with specific design features of the planning and accountability systems. In the New York SIP program, for example, the role of a liaison was created who provided the bridge between schools and the external project. The liaison initially spent five days a week at a site to help a school to focus on self-examination and development. The above cited program evaluations pointed to the crucial role these liaisons played for the success of planning. Another example is a system installed in Australia and described by Cuttance. In this system instances of external review of externally imposed performance goals alternate with instances of internal review of school development (Cuttance, 1994).

We are now in the position to specify our questions. We saw how earlier education reform policies often failed to affect a school's instructional program and how earlier planning efforts often failed to focus schools on key tasks of improvement. With new comprehensive accountability systems in place in many jurisdictions we want to see if this new generation of policy designs does a better job in focusing the school and in shaping the schools' instructional programs. Presumably, the more comprehensive and coherent character of the new accountability systems might make policy levers more forceful. Moreover, in studying schools on probation we

selected a group of schools that, because of the threat of sanctions imposed on them, should experience a particularly strong incentive to comply with external demands. But on the other hand, these are also the very schools where, because of performance history and lack of capacity, internalization of these external demands could be particularly burdensome. Thus, our analysis of school improvement plans will show whether we can find documentary evidence for goal focus, programmatic alignment, and internalization in the schools' improvement plans. Since our analysis for this paper is limited to documentary analysis, the issue of internalization can only be addressed by searching for documentary traces of presumed internalization processes. Data from actual case studies, furnished at a later date, will give us a more complete picture. Lastly we want to see if policy design differences make a difference in the way schools focus, align their program, and internalize external demands by interpreting their task in light of site-specific conditions for change.

The Three Systems

We investigate three accountability systems located in the states of Maryland and Kentucky and the city of San Francisco. At the time of our study, all three jurisdictions had been engaged in school accountability for some time, and all three had created fairly comprehensive systems consisting of performance indicators, sanctions, and formalized treatments for schools performing at various levels. All three had created a category for schools on probation. We will not describe the three systems in detail here, but highlight two features of the systems that are most relevant for our analysis. We will pay primary attention to the way the systems measure performance and address the bridging of external demands with internal development.

Maryland

The Maryland school accountability system has been in place since 1993 and, unlike systems in other jurisdictions, has been fairly stable over the last five years. The state created its own performance-based test that it administers to all elementary and middle schools. High schools are slated to give a performance-based test at a later date. To reduce variation, our study only deals with elementary and middle schools, i.e. with the types of schools whose performance evaluation depends to a large degree on how well their students do on the performance-based test (MSPAP = Maryland State Performance Assessment Program). In addition, Maryland measures schools' performance based on how well their students do on a basic skills test (MFT = Maryland Functional Test) and based on attendance rates. Students are tested in third, fifth, and eighth grades. The state has composed a formula that establishes a school's performance index. This formula attaches different weighs to the three performance indicators, MSPAP, MFT, and attendance rates. [1] Accounting for seventy percent of a school's performance index, the heart of the Maryland accountability system for elementary and middle schools is the MSPAP that assesses students' ability to perform a broad range of complex activities with often real-life applications in reading, writing, mathematics, science, and social studies. The state department of education has established satisfactory performance benchmarks that all schools are to reach

---

[1] Where *DS* = Distance from Satisfactory, for elementary schools:

$$SPI = \frac{Attendance\ DS + MSPAP\ Grade\ 3\ DS + MSPAP\ Grade\ 5\ DS}{13}$$

The *SPI* for middle schools adds Maryland Functional Test (*MFT)* scores:

$$SPI = \frac{Attendance\ DS + MSPAP\ Grade\ 8\ DS + MFT\ Middle\ DS}{8}$$

$$\text{Distance from Satisfactory} = \frac{School's\ Performance\ (Numerator)}{Satisfactory\ Standard\ (Denominator)}$$

within a period of ten years or so. For the MSPAP, for example, that benchmark is a rate of 70% of a school's student body passing the test at the "satisfactory" level. Schools gauge their performance expectations and establish improvement goals according to the state's performance benchmarks.

In Maryland, the state superintendent and board of education designate schools as "reconstitution-eligible," i.e. on probation, that are performing in the bottom rank and have had declining performance records in previous years (Maryland State Department of Education 1997). Upon identification as "reconstitution-eligible," the school is required to submit a school improvement plan for the state's approval. Review panels and state monitors visit the school to help in the diagnosis of needs and oversee implementation of the plan. The state monitors, however, do not have the role of change agents in schools. Though often experienced educators or retired administrators, they have neither received extensive training for their role, nor do they have the time to get intensively involved in individual schools. Funding and organization of capacity-building measures for individual school sites are largely left to local districts which in the case of Maryland tend to be large and congruent with counties. State-provided funds vary from year to year, and per-school allocation varies with the numbers of schools in the program.

Up to Oct. 1998, the state put 89 schools on probation. Most of these schools perform in the bottom rank with recently declining test scores and are faced with high educational loads. Of the 82 elementary and middle schools of these schools-on-probation, students qualifying for free and reduced price lunch range from 32% to 100%, with a median of 77%. Most of the "reconstitution-eligible schools" (RE schools) in Maryland serve an African-American student population (over 80% of the enrollment in most RE schools); the overwhelming majority (five sixths) of RE schools are located in the state's largest city, 10% are located in another local

jurisdiction with a majority African-American population, and two schools are located in other counties. For the 1996 cohort of RE-schools, mean percentage of students passing the MSPAP with satisfactory performance are 8.9% in math and 8.7% in reading (RE elementary schools), and 10.9% in math and 7.6% in reading (RE middle schools), a far cry from the state's 70% benchmark. Up to this point, none of the schools has successfully exited the system, nor have final sanctions been applied to any of them thus far. "Reconstitution-eligibility" in Maryland, rather than a transitory stage, appears to be a protracted period of probation for schools facing exceptionally arduous challenges. In summary, the state measures schools' performance with a small number of purely quantitative and clear indicators that center around student achievement on ambitious performance-based tests. An explicit bridging feature is not part of the design. The state monitors' role is primarily external oversight. Capacity building measures are left up to local districts. Probationary schools in Maryland are bottom performers.

Kentucky

In 1990, the Kentucky General Assembly passed HB 940, The Kentucky Education Reform Act (KERA), in response to a 1989 Supreme Court decision which declared the Commonwealth's system of public schooling to be unconstitutional due to the inequity and inadequacy of funding provided for schools. KERA created a comprehensive system of governance and accountability strategies intended to encourage all students to perform at high levels. The Act created accountability standards for all students, provided curriculum content guidelines and mandated assessments that reflected these standards. KERA also required that schools be governed by a site-based decision making (SBDM) council comprised of a body representative of parents, teachers and administrators. In addition, KERA provided a statewide

system of professional development and also devoted considerable funding to the creation and maintenance of a technology network. In addition, the General Assembly also set up a funding method for providing equitable financial support to all local districts.

KERA outlined six learner goals specifying what all students should know and be able to do. Progress towards these goals was to be assessed through the Kentucky Instructional Results Information System (KIRIS), recently redesigned and renamed CATS (Commonwealth Accountability Testing System). The test results include student responses on open-ended response and multiple choice questions on the formal assessment of reading, math, science, social studies, art/humanities and practical living/vocational courses, as well as scores on student writing portfolios. These academic components combined with non-academic data such as measures of attendance, retention, dropout rates and the successful transition to adult life results in a composite index score for each school. Based on this score, the Kentucky Department of Education (KDE) sets a school-specific baseline index, and biennial school-specific growth targets. This is the level of achievement for which the school is held accountable. According to this system, schools are held accountable for continuous improvement of their scores. Schools are not compared to other schools, but all schools are to reach the same achievement level after a period of about twenty years regardless of initial baseline performance. Schools surpassing their predetermined index are eligible for rewards, while those schools falling below the mark are eligible for state assistance in efforts to improve the school.

Schools designated as "in decline" or "in crisis" were required to participate in the School Transformation and Renewal Program (STAR). Once in the program, they were required to write a school transformation plan and were assigned a Distinguished Educator (DE) to assist in the implementation of the Transformation Plan. Recently, the system has been changed. Initially,

intervention in schools not performing at expected levels was mandatory. However, since 1998 schools "in-decline" may opt out of the state assistance program, but they are still required to write and submit a plan.

As part of STAR, Kentucky created the role of the Distinguished Educator (DE) or recently renamed, the Highly Skilled Educator (HSE). This program component was designed to assist those schools performing below their baseline on the assessment with the help they needed to meet their accountability threshold in the next biennium. DEs or HSEs are teachers and administrators who work with schools to improve their curriculum and instruction and to implement the school improvement plan. With regards to school improvement planning, the HSEs lead schools through the planning process by assisting in the collection and analysis of data, identification of causes of decline, and provision of feedback about the reasonableness of goals and suggestions of possible strategies for reaching these goals.

Up until 1998, 153 schools have been put on probation. Forty-nine schools exited the program after one biennium, i.e. they met their growth target. Only eight schools have been on probation during two consecutive biennia.

San Francisco[2]

The ancestry of the San Francisco accountability system is not rooted in the system's productivity per se, but in efforts to desegregate the district's schools and to make education more equitable for its African-American and Hispanic student population. In describing the system, we will emphasize features that were in place for the period during which we collected

---

[2] The description of the San Francisco system follows a paper by Goldstein, Kelemen& Koski (1998)

school improvement plans. The system underwent design changes prior to and after this period.

The purpose of the Comprehensive School Improvement Project (CSIP), begun in 1993, is to target underperforming schools that serve large numbers of African-American and Hispanic student populations. Equity and productivity goals are spelled out in the philosophy of CSIP and in Board of Education sanctioned educational goals. These goals as well as the philosophical tenets attached to the policy are binding for schools in CSIP. The district evaluates the performance of all schools based on a fairly extensive list of quantitative and qualitative indicators. The quantitative indicators measure students' performance on the California Test of Basic Skills, truancy, attendance, suspensions, student drop-outs, transfers, funds, and staffing patterns. In addition, qualitative indicators are used by an on-site review committee to evaluate the school. This committee of high ranking district officers visits the school once for one hour of inspection and once for a presentation by the staff during which the staff demonstrates progress towards the accomplishment of board-sanctioned goals. These goals broadly touch upon student achievement, parent and community participation, safe learning environments, and an integrated instructional and social support program for students. During the visitation, the committee reviews the school's improvement plan and a portfolio that demonstrates the school's progress, particularly with respect to the education of African-American and Hispanic students. Thus evaluation is at once rather comprehensive and arbitrary since the school cannot predict the committee's dictum.

Based on this intricate evaluation scheme, the superintendent and his staff identified up to nine bottom-ranked schools per year for CSIP, i.e. these schools are put on probation. Schools that failed to improve once in CSIP were actually reconstituted, i.e. building staff were forced to vacate the building. Upon entering CSIP, the school receives additional discretionary funds and

the service of a liaison, a district administrator who works with the site two days a week. Between 1993 and 1998, twenty-four schools entered the program. Of those twenty-four schools, ten were reconstituted. With the exception of two schools that are still in the program, the rest showed sufficient improvement to leave the program. However, the district does not make the actual performance ratings available to the public so that no clear statements as to the effect of the program can be made at this point. In summary, in San Francisco school performance is evaluated through a fairly extensive number of quantitative and qualitative indicators. Interpretation of these indicators and decisions on probation and reconstitution are left to the discretion of top district administrators. External demands and internal development needs are bridged with a district liaison. Schools' improvement efforts are to be tightly aligned with district goals, objectives, and philosophy centering on both tenets of effective schools and social and equity concerns while the scope of actual performance evaluation is fairly broad and diffuse.

The Designs in Comparison

We compare the three probation designs based on two criteria: the way the systems evaluate or measure school performance and the way they address the problematic of internalization (see Table I). With regard to the former, both Maryland and Kentucky have a fairly limited number of quantitative indicators that center on student achievement. The core of performance assessment in both states is a state-wide test that stresses higher-order thinking. In both cases, the test requires teachers to emphasize a form of pedagogy that fosters complex writing, reasoning skills, and student-generated meanings. In many instances, successful mastery of the test presumably forces teachers to learn new instructional skills and technologies. In San Francisco, performance indicators are broad in scope. The CTBS, the quantitative student

achievement test used by this jurisdiction, is one among many indicators applied to schools and does not require the same kind of instructional innovation that the two state tests entail.

Schools' responsiveness to external accountability is indicated by a correspondence between system design features and patterns found in school improvement plans. We hypothesize that if schools are responsive to the performance indicators of the accountability system we would find different patterns of activities across the three jurisdictions documented in the school improvement plans. In the two states, responsiveness would be indicated by schools focusing their needs analysis and activities on matters of curriculum and instruction. Activities would speak to new forms of pedagogy and instructional technology demanded by the performance-based test. In San Francisco, the scope of planned activities would be broader and include more strongly social and community components. Perhaps, the overall number of activities, suggested in the school improvement plans, would be higher in a system such as that in San Francisco, since the larger number of performance indicators would compel schools to cover more territory in their plans.

Our second comparative criterion has to do with design features that facilitate internalization. Across all three jurisdictions, internalization of external accountability could be indicated by a number of characteristics: by the formulation of a school mission and philosophy that profiles the school and serves as a filter (Conley, 1993) to screen out an abundance of external demands and expectations; by the search for those causes of under-performance that are internally controllable; and by external goals that are reinterpreted in light of actual school improvement capacity.

With regard to the designs examined here, we classify Maryland as having no specific design features in place that would bridge external accountability with internal development

needs and Kentucky as bridging the external/internal divide with the assignment of a specially trained change agent to schools on probation. It is conceivable that the strong presence of a "distinguished educator" or "highly skilled educator" in schools on probation helps schools in the task of self-examination and encourages them to engage in change activities that get closer to the core of teaching and learning. In the case of Kentucky school improvement plans, as compared to Maryland, this would be reflected in a stronger inward-looking gaze in the analysis of causes of under-performance and a more frequent mentioning of instructional change activities that are more closely related to teachers' core routines. More than the other two jurisdictions, the San Francisco design should enable schools to plan according to local conditions since performance indicators are more holistic and schools can demonstrate accomplishments through portfolio reviews. Yet, concentration on local school needs may be offset by a lack of goal focus and superficial comprehensiveness.

Table I **Three Accountability System Designs**

	Maryland	Kentucky	San Francisco
Performance indicators	Achievement (MSPAP, MFT), Attendance *all quantitative*	Achievement (CATS previously KIRIS) Attendance, Retention, Dropout rates, Successful transition to adult life *all quantitative*	Achievement (CTBS) Attendance Drop-out Transfers Suspensions Truancy referrals + School improvement plan School portfolio Presentation to district committee, *quantitative and qualitative*
External/internal bridge	State monitor (external supervision) Local assistance	"Distinguished Educator" "Highly Skilled Educator" (external supervision/ internal change agent)	District administrator liaison Local evaluation committee

In our comparative content analysis of plans from the three systems we will search for traces of the described design differences in school improvement plan patterns. But before we proceed to this analysis, we need to look more closely at the templates that schools are required

to follow when they engage in the writing of the plans.

The Planning Templates

Much of what becomes documented in school improvement plans depends on the templates schools are given for writing the plan. In all likelihood, schools will focus their thinking on aspects they are specifically required to cover. Those will also be the ones on which we will pick up in content analysis. Neglected areas or silences in the plans, on the other hand, could either reflect the externally imposed templates or schools' own thinking, in all likelihood the former, rather than the latter. Thus, the templates establish the limits of our document-based analysis of schools' dealing with planning. Overall, planning templates are similar across the three jurisdictions following the contours of the basic rational model of goals, needs assessment and action plans. But important differences among the three jurisdictions should be noted.

The *Maryland* SIP template is geared towards inducing a model of school improvement that is standards-based and combines managerial and cultural aspects of the process. In the typical Maryland school improvement plan, schools are to start from an analysis of needs and causes that lead to under-performance, using the diagnostic tools of the accountability agency. Goal formation consists of a section on the school's philosophy divided into vision and mission and a section on goals and objectives. Needs analysis and goals are to result in suggested strategies for improvement that are enumerated in an action plan. The action plan also identifies individuals or groups at the school site that are held responsible for implementation. Ideally the plans should be internally consistent, i.e. needs, causes, goals, and activities should align and rationally follow from each other. Presumably, then, the plans would facilitate the schools' transformation from a pattern of inefficiency to one of rational planning and management of

change that rests on an affirmation of schools' core beliefs and culture.

Schools wrote the school improvement plans according to the state-required format. Local districts assisted schools in writing the plans according to the state template. On the average the plans we analyzed are two-hundred to three-hundred pages thick. The state altered the format in 1998 resulting in plans that are below one-hundred pages long. Most of the plans analyzed for this paper are of the longer version, giving us insight into the way schools think about their improvement efforts long hand -- though filtered through the state-mandated template.

In *Kentucky,* school improvement planning has undergone dramatic changes since the inception of KERA. Beginning in 1992, all schools-on-probation in Kentucky were required to complete an annual School Transformation Plan (STP) which outlined their needs, goals and strategies for improvement during the coming year. In 1997, the Kentucky Department of Education redesigned the process of school improvement planning by "consolidating" the multiple plans needed for federal and state programs into one form to be completed every two years. This Consolidated Plan (CP) replaced the STP as the method of school planning throughout the state, and each school, not only those identified by the accountability system, were required to complete the process. Both systems of planning focused on instructional improvement and student achievement, but consolidated planning serves as the school's funding application for numerous federal and state funds.

The plans begin with the needs assessment component which leads schools through an evaluation of numerous aspects of their school. More than one hundred specific, in-depth, data-based questions are included in this assessment of the school. These areas include assessment scores, content area sub-scores and scores based on student group divisions, curriculum, student

grades, attendance, discipline, retention, transition into adult life, student health and well being, instructional support programs, special education programs, instructional materials, technology, professional development, school climate, parent involvement, and community support. This needs analysis is intended to guide the remainder of the planning process. The plan itself is comprised of two main sections, the executive summary and action components. The executive summary includes a mission statement as well as a narrative section which outlines the processes and persons involved in the creation of the plan. The specified outline reveals the logic behind the planning process and the expected outcomes of implementation. In addition, this section focuses on how the plan will be reviewed, opportunities for public comment and response and methods for sharing the plan and other information with others. Action components follow the executive summary. For each action component, the school must identify a priority need, a goal which addresses the need, causes or contributing factors based on the needs analysis, and measurable objectives which address the causes or contributing factors. Following this are the school's strategies for meeting the objectives and goals. For each strategy listed, several additional items must be included: the expected impact, how this impact might be measured, persons responsible, start and end dates, and estimated costs and sources of funding.

The *San Francisco* improvement plan template also reflects a prescribed district-wide format. The plans are much less extensive than the Maryland or Kentucky plans, consisting of 10-15 pages as compared to 100-200 pages. They are less analytical and do not include narratives. The San Francisco site plans begin with an "analysis of current conditions." By that is meant a list of the district and superintendent's goals and areas of concern and a rating of high, medium, or low regarding the school's needs in a given area. Vision statement and site priorities follow, as well as lists of planned activities. The activities are matched to particular goals and

areas of concern. School background information is provided by a separate document called School Accountability Report Card (SARC).

Sample

Our documentary analysis from *Maryland* is based on 46 school improvement plans that were compiled by "reconstitution-eligible" schools in the state of Maryland. The forty-six schools, representing a little over half of all RE schools, were selected by balancing three criteria. The sample should reflect the proportion of the reconstitution cohort, the distribution of school types (elementary, middle, high) in the universe of all RE schools, and the local jurisdictions involved. As was mentioned above, only elementary and middle schools are included.

Table II **Maryland Sample Selection by Cohort, Type, and Local Jurisdiction**

Cohort	Total number of plans	Elementary schools	Middle schools
1995	3	1	2
1996	16	12	4
1997	9	6	3
1998	(Local A) 9	4	5
1998	(Local B) 9	6	3
Total	46	29	17

Kentucky plans were selected from the universe of schools enrolled in the STAR program according to the following criteria: by region, by status in the STAR program, by school type, and by planning template, adjusted to a sample of about forty plans. This resulted in a total of thirty-seven plans, of thirty-one are elementary and middle schools, the focus of our study.

Table III **Kentucky Sample Selection**

Biennium	Total number of plans	Elementary schools	Middle / Junior High schools	High schools
1996-1997 (STPs)	11	3	5	3
1998-2000 (CPs)	20	6	12	2
Total	31	9	17	5

For *San Francisco*, plans for all elementary and middle schools in CSIP were read. This amounted to twenty out of a total of twenty four CSIP schools. The sample includes sixteen elementary and four middle schools. The high school site plans were not included for consistency among the three jurisdictions.

Content Analysis

A total of 103 plans from the three jurisdictions were read, coded, and rated with the help of a code book. Our analysis is based on 93 plans. The code domains followed the format of the plan: needs, causes, diagnostics, philosophy, goals, activities, resources, and responsibilities.

Activities were subdivided into organizational, climate, parent and community, curriculum and instruction, and professional development activities. A separate activity sub-domain *teacher performance* specifically hones in on new requirements for teachers' work performance (e.g., lesson plans, evaluations, etc.). The data base contains other domains (e.g., external partners, material) on which this paper will not report. A data base manager (MS-Access) was used to structure the coding. To facilitate coding and subsequent analysis, the data base contained a drop-down list of potential entries (i.e. school needs, tenets, goals, and activities) that the plans might contain. These lists were gleaned from the initial holistic reading of a smaller number of school improvement plans. Content that could not adequately be represented by one of the specifically listed entries on the drop-down lists were written into the data base. Entries were then categorized. For example, in the *cause domain* causes were categorized according to "internal" or "external" attribution of the addressed problem; in the *organizational activity* domain according to "new specialized service," or "whole school overhaul." In addition to these lists, readers were asked to rate activity entries with respect to priority, time frame, alignment, and responsibility. The coding concluded with a number of summary ratings with respect to the plan's focus, alignment, and consistency.

The plans were read independently by trained readers, two for each jurisdiction, who shared the reading load. A small number of plans was initially read by the whole research team. Agreement on codes and ratings was established. Coding reliability for the Maryland and Kentucky plans was established through the following procedure. Twenty of the 46 Maryland plans were read by two readers at various intervals (i.e. ten from each reader overlap). We compared the codes and ratings of these twenty plans with each other, as well as all plans read by one rater with all plans read by the other. Presentation of the findings is based on the forty-six

coded plans that are "first reads." Reliability was checked in two ways: first, all plans read by one reader were compared with all read by the other; the two data sets overlap with twenty of the forty-six plans. If similar values obtained between the two data sets, ratings and codings were deemed reliable. This was done for each analyzed rating or coding. Second, when values diverged in this procedure, only the plans that were read twice were matched and similarities were checked. In this way, we checked reliability for all plans and those that were read twice. For the Kentucky plans, ten of the thirty-seven plans were read twice. Ratings for these plans were compared. For San Francisco, only two of the twenty plans have been read twice up to this point and ratings were compared. As a result, San Francisco findings are presented with less confidence. Codings and ratings that were deemed unreliable (e.g., some of the summary ratings) were deleted from the analysis. In this paper, we only report on findings for which we could obtain agreement between the two raters within a margin of no more than plus or minus five points. We will report inter-rater reliability scores whenever greater deviations occur.

The content analysis is part of a larger research project, funded by the Office of Educational Research and Improvement, on the effect of reconstitution or probation on school improvement.[3] The analysis of school improvement plans is complemented by case studies in at least 12 schools.

Findings

In presenting findings from the content analysis, we will first analyze patterns we find in the school improvement plans for each jurisdiction separately. In a second step, we will look at

---

[3] We thank OERI (grant # R308F70035) for its generous support of this research.

patterns across the three jurisdictions. Finally we will draw inferences from correspondences between identified patterns and specific design features. Following the templates of the school improvement plans, we will look at the following components: analysis of needs, causes of decline, goals and philosophy, and action plans.

As a measure of focus and alignment, we want to know to what degree a school's needs analysis corresponds to the performance indicators of the accountability system. In looking at what sort of causes schools enumerate for their decline or insufficient performance, we were particularly interested in finding out whether schools approached their problems with a sense of efficacy and control (Ashton & Webb, 1986; Weiner, 1986). We tried to capture this by identifying causes that were enumerated in the plans and by classifying whether these causes were related to external factors in the school's environment or to internal factors under the direct control of educators at the site. We assumed that schools which focus their attention primarily on factors under their direct control treat school improvement efforts with more efficacy and demonstrate a more internalized model of change.

One often documented problem with ineffective schools is their lack of clear goals. Accountability systems potentially streamline the organization's thinking by aligning school goals with the performance expectations of the accountability agency. When embedded in an accountability system, goal formation entails two operations for schools: (1) goals ought to reflect the expectations of the accountability agency — expectations that will undoubtedly be high for most schools on probation; (2) goals ought to reflect the site conditions of the school, i.e. they should represent realistic growth goals in order to inform and guide internal steps of improvement. Thus, when schools formulate goals they negotiate the tension between high external expectations and what practitioners themselves deem realistic. In analyzing the *goals*

domain, we want to know whether schools are able to formulate clear goals, and how these goals solve the tension between external expectation and internal practitioner knowledge of what can realistically be achieved.

All school improvement plans contain a section on vision/mission/philosophy, here summarily referred to as the domain of philosophy. In this section, schools have the opportunity to enumerate the basic tenets that guide their work. They also have the opportunity to profile core beliefs and key ideas that guide their improvement effort for the year. Making inferences from statements of philosophy about a school's change process is difficult. What might be grandstanding and a collection of clichés in one case may represent the hard labor of internal reflection in another. But it could be useful to know whether schools use this section as a filter or as a means to profile a moral and cultural focus or whether an approach of "covering all bases of conventional wisdom" prevails as was found in previous planning attempts. The latter would undergird the external legitimacy function of the plans.

In looking at action plans, we want to know what kinds of activities schools consider powerful in achieving their improvement goals, whether stringent performance accountability made the schools focus their strategies, and to what degree these activities reach the core of instructional routines, an area that has been said to be particularly hard to reach. For all investigated dimensions, we consider how identified patterns speak to concerns of focus, alignment, and internalization. Lastly, we will consider how these patterns play out across the three jurisdictions, and we will try to construct correspondences between differences in focus, alignment, and internalization and specific design features of the three accountability systems.

Maryland

Naturally, documented needs must reflect the priorities of the accountability agency somewhat, or else the plans would not have gained official approval. But it is conceivable that schools adapt these priorities to specific site conditions, making the needs analysis more meaningful to the local school. In the Maryland case, student achievement, attendance, and climate (student discipline) are almost exclusively mentioned as needs in schools' analyses. Not surprisingly, all 46 plans mention low student achievement, as measured by the state's performance assessment, and negative climate measures. About three quarters of the plans document improvement of attendance as a need. By comparison, needs that are not directly measured by the accountability system are featured less. For example, lack of parental involvement is mentioned by only 11% of the schools (13% climate). Thus, the needs analysis of all 46 schools is aligned with the performance indicators of the accountability system. By comparison, needs not officially measured appear as less significant in the plans.

In the Maryland plans, about seventy percent of all *causes of decline* mentioned can be attributed to external factors. Interestingly, this somewhat conflicts with the plans' needs analysis which de-emphasizes parents and community. Typical external attributions for problems include scarce resources, high student mobility, and low socioeconomic environment. Thirty percent of the causes mentioned in the 46 school improvement plans are attributed to internal causes and thus directly controllable by the school's efforts. Schools highlight as internal causes shortcomings of specific teacher groups or specific individuals, organizational-structural issues, limitations in teachers' skills and knowledge, and leadership weaknesses. Although the externalization of problems may make eminent sense for schools that are socially beleaguered as some of these schools on probation are, one could have expected more emphasis on school-

internal causes of decline. That is — if schools consider the task of improvement as an internal project for which educators feel foremost responsible.

In most Maryland SIPs, the goals flow from the quantitative performance indicators set by the accountability system. Schools almost exclusively focus on improved achievement scores on both the performance-based and basic skills tests (MSPAP and MFT), attendance rates, and school climate measures (e.g., office referrals, suspensions, measures on a standardized climate survey). These goals account for 91% of all entries in this domain.

Numeric goals are gauged to the expectation of the accountability agency. In the case of the performance-based test which is by far the most important measure of school performance in the state's weighted statistical assessment formula, this translates into schools reaching the proclaimed threshold of 70% of students passing at a "satisfactory" level within a fairly short period of time, usually between three to five years. The discrepancy between actual school performance and external performance benchmarks is stark across most of the forty-six schools. According to our calculations on 33 reconstitution-eligible schools (1996 cohort), schools on probation in the state of Maryland increased the number of students scoring at the "satisfactory" level by an average of 3 to 5 percent over a two-year period (depending on subject). But promised gains often reach two-digit figures. We cannot judge from the SIP document whether awareness of these discrepancies was in fact generated by reflecting on external goals, we can, however, infer that internalization is shallow when schools calculate their performance goals according to an external formula without addressing the gulf between projected high growth and past performance record. We found in 90% of the 46 plans this formulaic procedure for goal formation. We concluded therefore that in most schools goal formation happens in "conformity to system expectations" with "achievability doubtful"(code).

Our data indicate that a "covering all bases" approach prevails in the philosophy domains of the 46 coded Maryland plans. On the average, 4 tenets were coded per school. Twenty-eight of the 46 schools mention as a tenet (in one formulation or another) that "all students can learn." Twenty-four schools mention tenets circumscribed by the code "Preparation for the competitive society/technological age" and the code "Development of individual students to their fullest potential." Other statements prevalent in the philosophy domain are: "High expectations for our students" (20 schools); "Safe environment" (19 schools); "School as family/community" (17 schools); and "Contributing to democratic and diverse society" (15 schools). These tenets make up 85% of all coded tenets. Overall, the 46 schools do not use the philosophy section to profile their organizational uniqueness or to document moral decisions, for example a decision for or against tracking; rather they cover a swath of at times contradictory tenets that have high symbolic currency in the present reform movement. The problem with the "covering all bases" approach is not that it is invalid — it may actually be an approach that makes most sense to schools given the array of competing demands placed on them. The problem is that such philosophies do not help schools sort through the enormous fragmentation that these competing demands create.

Activities were coded according to seven domains: organization, governance and coordination of action, climate and attendance, parents and community, curriculum and instruction, professional development, and (expressly) teacher performance. The analyzed plans attest to an extensive array of activities. Table IV displays the total number of activities for all schools, the range, median for each school, and the totals per domain.

Table IV **Distribution of Activities by Domain (Maryland)**

ACTIVITIES DOMAINS	Total Activities (based on 46 schools)	Median (per school)	Range (activities per school)
Organization	200 (9%)	4	1-9
Governance	132 (6%)	3	1-7
Climate and attendance	354 (17%)	7	2-17
Parents and community	226 (11%)	5	1-14
Curriculum and instr.	536 (25%)	11	3-21
Prof. development	456 (22%)	9	4-25
Teacher performance	209 (10%)	4	2-9
Total entries	2113	46	

As Table IV shows, the number of activities (2113) that the universe of all 46 schools attempts to undertake in one year is staggering. A set of close to fifty activities on the average amounts to a substantial reform load for a school. A focus on curriculum and instruction and professional development activities (47%) is detectable, but not decidedly so. Activities related to instruction in classrooms are clearly most numerous. Instruction-related activities in the professional development domain are also the most numerous category (about 70%). But schools also attend to many other areas of their operation; some may indirectly benefit instruction in classrooms. It seems that issues of governance and coordinated action (e.g., shared decision making), so prominent in earlier restructuring efforts, play a lesser role in accountability-driven reform. Thus, the rough quantification of activities suggests that classroom instruction is a focus, but not decidedly so. Issues of governance and coordination find relatively little attention.

We present findings on the content of activities by displaying an action plan of a modal school that we hypothetically constructed according to quantitative proportions of activities (Table V). The modal school allocates a certain number of activity slots to each domain in

proportion to the universe of all 46 schools. Each slot is filled by those activities that are most often mentioned by all schools for a particular domain in the order of frequency. Hence Table V displays the most frequent activities in each domain to the extend that the domain is covered in all forty-six plans.

Table V **The Modal Action Plan for Schools on Probation (Maryland)**

Organization	Governance/ Coordination	Climate/ Attendance	Parent/ Community	Curriculum& Instruction	Professional Development	Teacher Performance
New personnel	School Improvement Team	Tightening attendance procedures	Parent workshops	Packaged instructional program	Dimensions of Learning	Requiring use of specific curricula or materials
After-school extended program	Monitoring activities	Awards (assemblies)	Parent nights	Computer-assisted instruction	MSPAP/ Maryland Learning Outcomes	Requiring lesson plans
New specialized role/service	Task forces and committees	School-wide discipline plan	Parent/ community newsletter	Planning test-specific activities	Performance-based instruction	Tightening principal supervision
New teams		Conflict resolution program	Regular home calls	Remediating specific curricular weaknesses	Cooperative learning strategies	Peer coaching
		Spirit days	Parent shared-decision-making	Portfolios	MARS (performance-based Math)	
		More parent-teacher conferences		Cooperative learning	STARS (performance-based Science)	
		Display of student work/		Interdisciplinary integrative units or lessons	City Curriculum implementation	
		[Strengthening of counseling department (tie)]		New test-specific performance-based lessons or units	Reading instruction	
				Developing list of test words	Cooperative discipline	
				Writing process	Writing Across the Curriculum	
				Projects		
				Writing prompts		

BEST COPY AVAILABLE

The content of the intended activities is closely related to most frequently mentioned needs which coincide with school performance indicators; intended organizational changes have to do with additional instruction, attendance, and student discipline; governance activities are directly related to the task of managing and implementing the plan; all parent activities are directly tied to presumably most pressing problems that are indicated by performance measures (absences, tardies, discipline) with fewer activities stressing positive reinforcement, such as rewards and character education; curriculum and instruction as well as professional development activities are centered on the new challenges of the performance-based test (e.g., test simulation exercises, test language, importance of writing, portfolios, student-centered instructional strategies), on new district-adopted curricula and new instructional program packages. A preponderance of curriculum/instruction and professional development activities is related to external levers and initiatives (test, instructional packages, standard staff development workshops). Relatively little time is spent on lesson planning and instructional experimentation.

Relatively few activities tackle the problem of teacher commitment and motivation. In the modal school, nearly two-thirds of activities coded under teacher performance by the raters consist of surveillance activities, i.e. requirements to use specific curriculum, instructional practices, lesson plans, and tightening principal supervision. Activities reflecting teacher commitment, such as mentoring novice instructors and peer coaching, comprise 17% of all teacher performance entries.

In sum, the analysis of intended activities that schools document in the school improvement plans uncovers that schools, for the most part, design their action plan around the measured performance indicators of the accountability system and the upgrading of curricular programs initiated by districts. Schools have aligned their action plans with the performance

assessment system and district initiatives in its aftermath. Yet, aligning activities with performance indicators still yields a substantial load of activities in most schools, perhaps even overload, despite the system's encouragement to focus, especially on matters of teaching and learning. Schools, it seems, have a tendency to cover the whole territory mapped out by the performance indicators. Hence one can find in all school improvement plans activities listed in all domains. This extensive and inclusive coverage suggests to us that schools in all likelihood want to make sure that they answer to all areas in which their performance is measured by the external accountability system.

In summary, as far as goal formation is concerned, the accountability system succeeded in focusing schools on those elements of school life that it deems important to measure. These elements refer to the model of a good school spelled out in effective schools research. Moreover, goals flow from analysis of needs. Thus, on the level of espoused views of school improvement, the plans exhibit focus and more rational alignment between needs analysis and goals. But there are troubling signs of superficial internalization. The schools' externally-directed perspective on causes for their troubles and their formulaic method of solving the tension between external expectations and internally established margins of improvement betray an attitude of compliance, rather than ownership. Mandated goals, however, that show little adaptation to site conditions are poor instruments to direct school improvement (McLaughlin, 1990). Teacher performance requirements betray a spirit of enforcement as schools promise to increase surveillance of teachers through increased principal supervision and the writing of daily lesson plans.

The analysis of intended activities reveals that schools on probation approach their task with an attitude of meeting external obligations and internal incrementalism. This approach leads, on one hand, to a streamlining of change designs and (intended) upgrading of teaching

skills by aligning activities to the performance indicators of the system and to curriculum upgrading initiatives of districts. In this way, the accountability system may have succeeded in making schools more effective, assuming that external performance indicators target, indeed, the most essential aspects of a school's performance. On the other hand, the approach taken by Maryland schools betrays a notion of change that emphasizes extension of existing programs (i.e. more instruction, more personnel, more skills) rather than internal cultural shifts, and extensive covering of performance "territory" rather than concentration on central themes for the year or intense work on routine lesson planning. The burden of responsibility carried by administration and special services, the neglect of teacher commitment activities, and the notion of professional growth as skill and workshop-bound all corroborate a pattern of change that couples external obligation with traditional incrementalism. Whether this strategy is adequate in terms of making schools more effective is not the main question of this paper. Here we are less interested in issues of effective strategy and more in issues of correspondence between planning pattern and policy design.

Kentucky

While we discussed the significance of analyzed patterns more extensively for the Maryland case, we will present findings for the other two jurisdictions in more brevity taking the analysis of the Maryland plans as our point of departure.

For schools in Kentucky, student test scores are the overwhelming focus of the accountability system, making this the primary area of concern for all schools. In response to this, each of the twenty-six Kentucky plans identified the issue of low student achievement on

the state assessment. Although very minor in comparison, only three schools cited low attendance as an area of concern. Although schools are accountable for attendance rates, this received minimal attention in school plans. It is interesting to note that more than two thirds of the schools identified negative school climate (lack of parental involvement) as a problem. Although parental involvement is addressed in the needs assessment process of the consolidated plan, it is in no way incorporated into the measure of school performance in the accountability system. Thus, there is a certain disconnect between performance indicators and identified needs that could be related either to a greater discretionary interpretation of the task on the part of schools in Kentucky (as opposed to Maryland) or to planning components of the STAR program that are brought to bear on the schools by distinguished educators. Overall, low student achievement and negative climate measures accounted for 93% of the needs identified within these Kentucky plans.

The schools represented by these plans offered numerous explanations for their decline. From the twenty-six plans, more than 140 causes were identified. Two thirds of the causes included in the plan were attributed to internal factors, predominately to organizational structure, neglect of specific skills, and limitations in teacher skills or knowledge. Due to the detail provided in the assessment reports and in the CP needs assessment, schools can pinpoint exact deficiencies in school programs, parent involvement and student achievement. The remainder of causes linked to external factors dealt with scarce resources, lack of parent support, student motivation, difficult students, and difficult socioeconomic status.

According to the accountability system, each of the schools cited as a goal increased student achievement. In a majority of the plans, schools not only identified the level of desired

increase in CATS composite scores, but also in specified subject areas. Although indirect, schools also targeted increased achievement through two other goals – strengthened curriculum and improved instructional strategies. In addition to increased achievement, schools also desired increased parental and community involvement. Again, although this is not measured directly by the accountability system, this goal is consistent with the identified needs of many schools in the sample. Other goals cited in the plans were improved school climate (10 schools), decreased discipline referrals (5 schools), increased staff confidence (4 schools) and improved attendance (4 schools). In the case of Kentucky, goals stated in the plans reiterate quantitative KIRIS/CATS goals, in addition to parental involvement goals. The readers of the plans tended to code the quantitative KIRIS/CATS goals as achievable in the allotted time frame. The readers pointed to schools' past performance (e.g., successful exit from the program).

The mission statement is a required component of the plan. As with the plans in Maryland, the mission statements of the plans included multiple philosophical tenets. But on average, schools included only two tenets into their mission. Half of the schools focused on individual student potential. Schools also included the creation of a democratic community (8 schools), high expectations for students (8 schools), that all members of the schools community assume responsibility for student learning (5 schools), an atmosphere of trust (4 schools), and the creation of a safe environment (3 schools). Surprisingly, the tenet *"All Students Can Learn,"* the major emphasis of KERA, was included by only 4 schools. Thus, as in the case of Maryland schools do not use the philosophy section to individualize their profile, though the message in the mission statements tend to be sparser.

Table VI **Distribution of Activities by Domains (Kentucky)**

ACTIVITIES DOMAINS	Total Activities (based on 46 schools)	Median (per school)	Range (activities per school)
Organization	86 (14%)	3	1-7
Governance	22 (4%)	1	1-4
Climate and Attendance	103 (17%)	4	1-9
Parents and Community	97 (16%)	4	1-10
Curriculum and Instruction	191 (31%)	8	1-12
Prof. development	95 (15%)	4	1-8
Teacher Performance	29 (5%)	1	1-3
Total entries	623	24	

. Table VI shows the distribution of activities across the specified domains. Overall these twenty-six schools identified 623 activities designed to result in improvement. Clearly, the schools are focused on curriculum and instruction and related staff development, as these activities encompass about 46% of all activities proposed, a percentage very similar to Maryland. Aside from curriculum and instruction, school activities were equally divided among climate, professional development, parent and community involvement and organization. The areas of governance and teacher performance were insignificant in the overall composition of the plans. The inclusion of strategies of changing the governance structure of the school were limited. In the twenty-six plans, only twenty-three activities were recorded. More than half of these suggested the use of internal monitors, a feature that is unique to the Kentucky planning process

and the decentralized management of schools in this state. Each section of the consolidated plan has a specified manager responsible for its implementation. In most cases, the plan required the component managers to report to the school's SBDM council (site-based decision making) at regular intervals on the progress of their components. This feature of internal monitoring and articulation with SBDM is absent in the other jurisdictions.

Table VII **The Modal Action Plan for Schools on Probation (Kentucky)**

Organization	Governance	Teacher Performance	Parent/ Community	Curriculum & Instruction	Professional Development	Climate
After-school/ extended day program	Internal monitors	Require use of specific curriculum	Increased communication with parents	Curriculum alignment	Instructional strategies	Awards/ Awards Assemblies
New specialized role/service			Volunteers	Remediating identified weaknesses	Best practices	Testing incentives
New personnel			Parent workshops	Portfolios	Teaching methods	Display of student work
			Parent newsletter	Writing prompts	Technology in instruction	Testing environment
				Technology		
				Changing basic lesson format		
				Adopting packaged programs		

BEST COPY AVAILABLE

Strategies aimed at modifying curriculum and instruction were the most utilized in the Kentucky school improvement plans. The most frequently mentioned activity was alignment of the curriculum. KERA mandates that all districts and schools receive a model curriculum framework which is directly tied to the goals outcomes and assessment strategies. These content guidelines assist schools to identify content related to the assessment. Schools model their curriculum to the content which in many instances results in a redesign of lessons and curricula

Second to curriculum alignment is remediating identified curricular weaknesses. Because the assessment results offer detailed sub-scores, schools can quickly determine specific skills or knowledge that was absent in the previous test group. This allows schools to address these weaknesses before the next test cycle. A great number of the remaining curricular strategies were also focused directly on the assessment such as the focus on writing portfolios, inclusion of writing prompts into daily instruction, drill and practice of test items, test-specific activities, lists of test words, focus on the writing process and writing across the curriculum. Aside from those curriculum and instruction activities focused on assessment, other strategies included the incorporation of technology in instruction, changing the basic lesson format, adopting packaged programs (principally Accelerated Reader), interdisciplinary units and multiple intelligence activities.

In summary, Kentucky plans exhibit patterns similar to Maryland plans. But in Kentucky, plans focus on fewer activities, needs analysis and goals are not as strongly aligned to the quantitative performance indicators of the accountability system, activities in the curriculum and instruction domain emphasize more internal curriculum development and alignment, schools stress more internal causes for under-performance, and quantitative goals are rated as within

reach of achievability. Overall, Kentucky plans more strongly exhibit patterns that speak to a more thorough internalization of the goals and procedures of the accountability system.

San Francisco

Of the 123 identified needs for school improvement among the twenty San Francisco plans, low achievement (32%) and negative climate measures (15%) comprised most of the entries followed by integration and equity, interagency collaboration, parental involvement, and program integration. Each of these entries reflects the mission published by the district. Unlike the Maryland and Kentucky plans, San Francisco plans report areas of needs for improvement, but they do not include a narrative analysis, only ratings of high, medium, and low need on the district and superintendent's "Goal and Area of Concern Chart." The most conspicuous difference between San Francisco and the two state systems is the lower incidence of "low achievement" as a documented need which indicates a more multi-faceted performance evaluation scheme that goes beyond productivity as the primary concern of school improvement.

As was described in the section on the three systems, goal formation for San Francisco schools-on-probation, i.e. schools in the CSIP program, differs from the two state systems in that in San Francisco Board of Education sanctioned goals and improvement objectives are standardized for all schools. Therefore all CSIP schools refer to the same goals and objectives uniformly in their school improvement plans although variation occurs because schools emphasize different objectives that speak to a school's specific shortcomings. The goals are broadly formulated, however, and the objectives are not very prescriptive, presumably leaving the school ample space to translate them into specific actions. Examples of objectives are: "To

improve teaching and learning to enhance the academic achievement of all students." or: "Continue to improve educational technology." There are no specific quantitative goals given to San Francisco schools, and none are enumerated in the plans. Thus, externally formulated goals and objectives are adopted by CSIP schools unchanged, but these goals are neither prescriptive nor do they clearly benchmark performance levels. Thus, it is up to schools to interpret external performance expectations in light of site conditions and putative sentiments of high-level district administrators. The loose connection between objectives and intended actions can be inferred from the plans in that all plans index a given intended activity with a list of numbers that refer to district objectives which the activity is presumably addressing. However, the actual substantive connection is not specified, giving the association between external goals and internal activity a rather arbitrary character.

In the philosophy domain, schools show more variation than plans in the other two jurisdictions. While half of the plans list a conglomerate of philosophical tenets that reflect the district mission statement, the other half of the plans illustrate mission statements unique to individual schools, for example, a mission reflecting the bi-cultural nature of the school's student population, focusing on equal opportunity, or coupling multi-culturalism and the arts. Some of the schools' missions succinctly highlight the individual school accountability report card, an exhibit of local autonomy not evident in the Maryland and Kentucky mission statements.

The quantitative distribution of activities by domain (see Table VIII) shows that the modal San Francisco school on probation plans more activities than its counterparts in either Maryland or Kentucky. San Francisco schools plan a median of 56 activities for the school year. Apart from that, the overall activity pattern of CSIP schools in San Francisco resembles that of

schools in the other two jurisdictions in terms of widespread coverage of all domains and a de-emphasis of issues of governance and coordinating action. Not unlike schools in the other two jurisdictions, San Francisco schools emphasize curriculum and instruction activities more than any other domain (37%), but this emphasis is not as pronounced as in the two states (close to 50%). Indeed, parent and climate activities together account for a block of activities that address social concerns of school improvement. These activities are not as well represented in the plans in the other jurisdictions. San Francisco's improvement philosophy, rooted in an ancestry of desegregation efforts, may be echoed here.

Table VIII **Distribution of Activities by Domain (San Francisco)**

Activity Domains	Total Activities (based on 20 schools)	Median (per school)	Range (activities per school)
Organization	113 (10%)	5	2 – 14
Governance	49 (4%)	2	0 – 5
Climate and attendance	211 (19%)	9	4 – 27
Parents and community	194 (18%)	9	6 – 13
Curriculum and instr.	282 (26%)	15	4 – 27
Prof. development	124 (11%)	6	3 – 11
Teacher performance	116 (11%)	5	1 – 15
Total entries	1089	52	

San Francisco's pattern of concern for social climate and equity is further corroborated when one looks at the actual activities our constructed modal school plans to undertake. Many

BEST COPY AVAILABLE

activities in the climate and parent domains have the purpose of strengthening social relationships between home, school, and wider community and are less instrumentalized for purposes of boosting the school's performance on discipline and attendance related matters. Likewise, the curriculum and instruction domain lists as its most prevalent activities *remediate identified curricular weaknesses* and *on-going experimentation with instruction.* Both of these accentuate more clearly internal development needs or interests of the schools.

Relative to Maryland, San Francisco school accountability is less of an enforcement scheme. While schools feature more strongly a requirement of colleagues to adopt specific curriculum and even to collaborate, notably absent are elements of increased surveillance, such as tightening of administrative supervision and the writing of daily lesson plans. This relative de-emphasis of administrative surveillance, together with elements of instructional experimentation, hints at a more teacher-centered change model in the case of San Francisco.

In summary, the San Francisco modal improvement plan refers to external objectives, but connections between objectives and activities are fairly loose. Needs are not clearly prioritized and philosophy statements are reflective of schools' uniqueness. The modal action plan covers a wide swath of activities and activity domains. It places a strong emphasis on social concerns of the school, yet instructional concerns are similarly important. A loose connection between district goals and activities notwithstanding, the plans reflect the philosophical thrust of the whole school improvement program. Prevalent activities in the curriculum and instruction domain reach the core of instructional routines and suggest a more internal impetus for instructional reform in San Francisco. A de-emphasis of surveillance activities bolsters a more teacher-directed approach to improvement.

Table IV **Modal Action Plan (San Francisco)**

Organization	Governance	Climate	Parent/ Community	Curriculum & Instruction	Professional Development	Teacher Performance
New specialized role/ service	New teams	Awards/awards assemblies	Parent workshops	Remediate identified curricular weaknesses	Literacy initiatives	Required professional development
Building renovation	Task forces/ committees	Multi-cultural assemblies	Increase parent representation	On-going experimentation with instruction	Reading instruction	Required use of specific curriculum
After-school/ extended programs		School-wide discipline	Parent/ community newsletter	Adopt packaged programs	Math instruction	Monitor student progress
Tutorials		Conflict resolution	Parent volunteers	Ambitious projects	Technology	Regular home contact
New personnel		Tighten attendance procedures	Parent nights	Portfolios	Conferences	Required ambitious projects
		Clean-up days	Parent meetings	Test-taking workshops	Visits	Required collaboration
		Display student work	Regular home calls	Combination of subjects		
		Service learning	Contact neighborhood associations	Interdisciplinary units or lessons		
		Mental health teams	Parent surveys	Curricular alignment		
		Deepen understanding of organizational philosophy	Parent conferences	Multiple Intelligences/ learning styles		
				Changing basic lesson format		
				Writing across the curriculum		
				Writing prompts		
				Cooperative learning/Writing process		

BEST COPY AVAILABLE

Cross-Cases Analysis

Earlier, we asked how patterns that we identified for each jurisdiction are similar to and different from each other with regard to focus, alignment, and internalization, and whether we can can construct correspondences between differences in focus, alignment, and internalization and specific design features of the three accountability systems. Table X displays a synopsis of our content analysis across the three systems.

Focus. Our analysis shows that schools on probation have a tendency to cover all domains of the organization, independent of design differences among the three accountability systems, though some noteworthy differences obtain. Kentucky schools are the most focused *relative to the other two jurisdictions,* as indicated by the lowest number of median activities listed, while that number is highest in San Francisco. Schools in all three jurisdictions emphasize activities in the curriculum and instruction domain. This is most markedly so in Maryland, less so in Kentucky, while in San Francisco social concerns addressed in the parent and climate domains attain equal, if not more weight. Thus schools in San Francisco are torn between equity and productivity concerns; in the Kentucky and Maryland accountability systems productivity weighs more heavily as far as documented views of schools are concerned. In sum, particularly the Kentucky and Maryland designs focus schools on instructional reform, but this does not stop schools from dispersing activities across all organizational domains.

Table X **Patterns in School Improvement Plans Across Three Accountability Systems**

	Maryland	Kentucky	San Francisco
Number of activities (median) per school	46	23	52
Primary activity domain	Curriculum and instruction	Curriculum and instruction	Parents and community and Climate
Primary activity within Curriculum & Instruction	Test practice District sponsored workshops	Curriculum alignment (lesson planning) Remediating identified weaknesses	Remediating curricular weaknesses On-going experimentation with instruction
Goals	External quantitative goals, achievability doubtful	External quantitative goals, school-specific growth targets	External goals and district-sanctioned objectives.
Causes of decline	70% external attribution	30% external attribution	n/a
Philosophy/mission	Generic, "conventional wisdom"	Generic, "conventional wisdom"	50% unique 50% aligned with district
Governance Fostering collegiality	Negligible	Negligible	Negligible

Alignment. External demands of the accountability agencies drive the writing of school improvement plans in all three jurisdictions. Externally generated goals are taken over unaltered by almost all schools. This could have been expected since a key component of external accountability systems is their insistence on "crystalline-clear" and pre-established goals and

objectives which limit school autonomy to the means of achieving these goals. But it is conspicuous that goals are almost never interpreted in view of site conditions and realistic growth expectations on the part of practitioners. A pattern of alignment obtains more strongly for the two state systems with their looming student achievement tests and stringent quantitative performance indicators than for the less prescriptive and directive objectives in the San Francisco case. In the case of Maryland, schools tailor many improvement activities to the test (writing, portfolio, test-taking skills) and to curricular and staff development initiatives put forth by districts to meet state demands. In Kentucky, a similar emphasis on new technologies of teaching, tied in with the test, is observable. But here relatively more emphasis is placed on internal curriculum alignment. In San Francisco, the array of suggested activities is broader, and instructional improvement is connected with experimentation more so than alignment.

Alignment of a different sort takes place independently of specific accountability system designs. In all three jurisdictions, educators hail extended services, assemblies, awards, and displaying student work, tightening attendance procedures, more enforcement of discipline, and increased parent involvement as a means to improve the organization. In the area of curriculum and instruction, schools deem, among others, the installment of packaged programs, portfolios, the writing process, test awareness, and interdisciplinary approaches as promising strategies.

How can these similarities be explained? Some of the highlighted activities are school routines, such as awards assemblies, that schools have traditionally used to influence school climate. While it is conceivable that they had fallen into disuse in schools-on-probation, they are certainly not innovative. Other highlighted activities speak directly to the orthodoxies of the effective schools movement, a model that has garnered powerful support among efficiency

oriented external accountability agencies. Clearly, the stress on test-taking skills is a common reflex of schools to the increasing urgency external agencies and the public place on tests. Lastly, ideas such as the writing process, portfolios, and interdisciplinarity have recently gained strong currency as legitimate innovations in professional circles. Thus, differences notwithstanding, our constructed modal schools across all three jurisdictions have in common an "alignment" of activities to routines, reflexes, legitimate models of good, i.e. effective schools, and legitimate innovations.

Following the theoretical models discussed above, we can presume that beleaguered schools on probation, independent of their local site and accountability context, highlight activities in their school improvement plans that are apt to capture support from the external environment because these activities are either uncontroversial and common-sensical routines or legitimate ideas and models that have strong currency with external accountability agencies or the profession. In that sense, school improvement is externally driven in all three jurisdictions, though there seems to be more room for experimentation in the case of San Francisco. Gaining external legitimacy may also be the engine behind a tendency of schools in all three jurisdictions to formulate missions and philosophies that conform to, what Conley in his analysis of school improvement plans characterized as, conventional wisdom, as opposed to using them as means to filter and profile.

Earlier we hypothesized that a critical issue in school planning under conditions of accountability is the relationship between external accountability and internal school development. We have strong indications that school improvement plans under these conditions answer more strongly to external demands than to internal needs of school development. To

summarize the various indications so far, schools adopt external goals on many instances ritualistically without analysis of their realism; their philosophy and mission statements tend to be boiler plate, conventional, and casting a wide net. Activities that could serve teacher commitment are de-emphasized; perhaps commitment is taken for granted, or those activities could give the "wrong" impression that educators spend precious resources on themselves, rather than on the betterment of services. Though the plans, particularly in the two state systems, are somewhat focused on curriculum and instruction, they tend to be comprehensive rather than strategic and speak to a strategy of covering all bases at once rather than of strategic or staged development, conforming to earlier observed unproductive patterns of externally induced planning (Levine and Leibert, 1987).

Yet, characteristic differences in the degree of internalization obtain in our data between the two state systems. The external gaze seems more strongly articulated in the Maryland school improvement plans. Characteristic differences show up in our data in the *curriculum and instruction* and *causes of decline* domains. As was pointed out above, the Maryland modal school tends more strongly to engage in curriculum and instruction activities that have to do with packaged programs, test awareness, and district initiatives. By contrast, Kentucky "schools in decline" feature more prominently school-internal curriculum alignment activities. Another remarkable contrast can be observed in the way schools analyze "causes of decline" in the two states. While Maryland schools locate about 70% of all mentioned causes in the external environment, Kentucky schools mention causes in their school improvement plans that were coded "external attribution" in only 30% of mentionings. Presumably, Kentucky schools, although under similar accountability pressures as in Maryland, approached their task with a

more inward-looking gaze.

How can we interpret these characteristic differences in focus, alignment, and internalization across the three jurisdictions? Is it possible to construct correspondences between patterns of improvement documented in the plans and design features of the accountability system? To begin with focus, it seems plausible that the clear quantitative indicators of the two state systems more stringently focus schools on curriculum and instruction. In San Francisco, the combination of a relatively larger number of quantitative and qualitative school performance indicators and the combination of equity and productivity concerns in the philosophical thrust of the policy may have dissipated focus. Nevertheless, in all three jurisdictions, school improvement plans tend to be comprehensive, rather than focusing on strategic next developmental steps or stages.

Likewise, programmatic alignment is more tangible in the two state systems with their reliance on a few indicators and a looming student achievement test that inheres novel instructional formats and technologies. By contrast, in San Francisco activities include more experimental elements in the instructional domain and more broadly targeted social climate activities. But despite a similar momentum for focus and alignment in both Maryland and Kentucky, the latter seems to have handled the bridging of external demands and internal needs differently. It is conceivable that in Kentucky the presence of a trained change agent at the school (distinguished educator, highly skilled educator), may have helped schools to treat external accountability demands as internal challenges, as evidenced by a stronger focus on internal causes of decline and curricular alignment activities that are more school-generated.

In summary, we conclude that both the nature of performance indicators and the way

bridging is addressed in the three accountability designs may have had an impact on the way schools plan their own change design. Clear quantitative measures centering on student achievement focus schools on issues of curriculum and instruction, and the services of an external change agent, who is specifically trained in the intricacies of the accountability system and is delegated to the school by the state, reinforces internalization of external accountability demands. Yet, these differences notwithstanding, the similarities in the plans across all three jurisdictions are striking. Perhaps the (smaller or larger) variety of performance indicators and the swath of correlates and philosophical tenets of the effective schools approach to educational reform was interpreted by schools to mean that school improvement should be commensurably comprehensive. Or it could be that a certain level of standardization in the business of school improvement has occurred through professional and managerial influences leading to the narrowing of innovative options that schools feel safe to chose from or are knowledgeable about.

If one follows our original discussion of patterns of planning in schools, it is likely that for schools writing a school improvement plan is in all likelihood, first and foremost, an exercise in shoring up external legitimacy, rather than a vehicle for internal development. This strong external orientation of schools can be tempered, it seems, by either more loosely structured performance indicators which open up more room for autonomous action on the part of schools, or by change agents who can bridge the external/ internal divide. Where, as in the case of Maryland, the system combines stringent performance measures with the absence of a bridging agent, the external orientation is most pronounced and we would surmise planning is least tied to actual school development events.

Limits of the Findings

In constructing a correspondence between a system's accountability design and schools' improvement plan patterns, we make assumptions without specifying the context in which accountability plays out in the three jurisdictions. We surmise that certain patterns could be associated with design differences, but we cannot conclude that with certainty. These patterns could be caused by other factors in the schools' environments that are not specified here. A few come to mind. For example, because of the way the three jurisdictions select their schools on probation, schools in Maryland are rock-bottom performers while schools in Kentucky are under-performers on various performance levels. It is conceivable that, on the whole, schools on probation in Maryland encounter more difficult socio-economic circumstances than Kentucky probationary schools so that Maryland schools might naturally look to their external environment to analyze "causes of decline." Thus we might see in this pattern an effect of school site, rather than one of accountability system design.

Another example, the overwhelming majority of "reconstitution-eligible" schools in Maryland are located in one large urban school district while the Kentucky schools are spread among many local districts. As a result, we might see in the Maryland patterns an effect of district, rather than accountability design. Moreover, with our formulation of *school* patterns we imply that the plans were actually written by the schools themselves when in fact they may been written by district consultants in Maryland or distinguished educators in Kentucky. Lastly, even in the best of circumstances, content analysis of documents alone can only uncover the espoused views of schools, and we might be hard pressed to interpret patterns in the plans without data on the context of implementation.

These are serious limitations and they compel us to read our findings more as hypotheses for further in-depth study, rather than final conclusions. In-depth study into the district and school site contexts is part of the wider study, even though neither one of the contexts is specified in this paper in any detail. But a few hints may be helpful. When we speak of schools' responses to accountability, as expressed in the plans, we should be more precise and speak of schools' and districts' responses because in both states, schools' responses are tinged by district intervention. Plans are official documents, subject to external review on the state level, hence they are of high concern for the external legitimacy of districts. In fact, in all our case study schools, plans were written by a faction of the faculty with external help. But even if we cannot make a clear distinction between district and school response, it is still significant to know what kind of plan passes muster in the different accountability systems. Hence, the correspondence between design features and document patterns reveals a shared conceptualization of school improvement and accountability among state officials who compose the design and districts and schools who interpret it. "Correspondence," then, refers less to direct cause (i.e. design) and effect (i.e. plan patterns) relationships, and more to a whole complex of interactions between an accountability agency and districts and schools on probation under its jurisdiction, of which design features may be just one component.

In any case, the analysis of documents for a larger number of cases provides direction for the more fine-grained study of implementation in a smaller number of cases. For example, if detected patterns from the documentary analysis hold up in these implementation studies, we would expect the plans' relevance in Kentucky to be higher for internal development than in Maryland. Moreover, we would expect to see more standardization in both Maryland and

Kentucky and more school site variation in San Francisco. Although we set out to learn from design differences, the findings from the content analysis are striking in their revelation of sameness across jurisdictions. The plans conjure up the image of schools on probation that approach planning with a externally directed concern for legitimacy or, if the plans are actually followed through, for compliance. In the plans, externally generated goals (of various prescriptiveness) appear as givens (Malen and Muncey, 1999); improvement efforts are comprehensive and tied to the effective schools model; and proposed "innovations" are the staples of the current consultant circuit. It will be interesting to see how this conceptualization of school improvement plays out in the schools struggling with improvement and to what degree it is relevant to explain their failures and successes.

A final word of caution is needed. We, in no way, tried to evaluate the three accountability designs. Since we began the study, accountability designs have evolved, districts and states have learned, and political priorities have shifted. Thus, we reflect on system designs at a particular moment in time. Interestingly, the Maryland and Kentucky designs have converged in some way: the plan templates have been streamlined in both jurisdictions; Kentucky made its strong external change agent feature optional for schools while Maryland has slowly moved in the direction of more technical support and bridging. In accountability research, it is up to us researchers to learn from the practitioners' and designers' cutting edge.

References

Ashton, P., & Webb, R. B. (1986). *Making a difference: Teachers' sense of efficacy and student achievement.* New York: Longman.

Ball, S. J. (1990). *The micro-politics of the school: Towards a theory of school organization.* London, New York: Routledge.

Bardach, E. (1986). Educational paperwork. In D. L. Kirp & D. N. Jensen (Eds.), *School days, rule days: The legalization and regulation of education* (pp. 124-44). Philadelphia: The Falmer Press.

Barth, R. S. (1990). *Improving schools from within.* San Francisco: Jossey-Bass.

Bell, T. H. (1974). *A performance accountability for system for school administrators.* West Nyack: Parker Publishing.

Biott, C., Easen, P., & Atkins, M. (1994). Written planning and school development: Biding time or making time. In D. H. Hargreaves & D. Hopkins (Eds.), *Development planning for school improvement* (pp. 80-90). New York: Cassell.

Broadhead, P., Cuckle, P., Hodgson, J., & Dunford, J. (1996). Improving primary schools through school development planning: Building a vision, exploring the reality. *Educational Management and Administration, 24*(3), 277-290.

Canner, J. (1985). *School improvement project: 1983-1984. Fifth annual process assessment.* New York: New York City Board of Education.

Clark, T. A., & McCarthy, D. P. (1983). School improvement in New York City: The evolution of a project. *Educational Researcher, 12*(4), 17-24.

Conley, D. T. (1992). *Strategic planning in America's schools: An exploratory study.* Paper presented at the American Educational Research Association, San Francisco.

Conley, D. T. (1993). *Strategic planning in practice: An analysis of purposes, goals, and*

procedures. Paper presented at the American Educational Research Association, Atlanta.

Constable, H. (1994). Three arenas of tension: Teachers' experience of participation in school development planning. In D. H. Hargreaves & D. Hopkins (Eds.), *Development planning for school improvement* (pp. 91-99). New York: Cassell.

Cuttance, P. (1994). The contribution of quality assistance reviews to development in school systems. In D. H. Hargreaves & D. Hopkins (Eds.), *Development planning for school improvement* (pp. 49-68). New York: Cassell.

Deal, T. E., & Peterson, K. D. (1999). *Shaping school culture: The heart of leadership*. San Francisco: Jossey-Bass.

DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review, 48*(April), 147-160.

Elmore, R. F. (1996). Getting to scale with good educational practice. *Harvard Educational Review, 66*(Spring), 1-26.

Firestone, W. A., & Corbett, H. D. (1988). Planned organizational change. In N. J. Boyan (Ed.), *Handbook of research on educational administration* (pp. 321-340). New York: Longman.

Firestone, W. A., Mayrowetz, D., & Fairman, J. (1997). *Rethinking high stakes: External obligation in assessment policy*. Paper presented at the annual conference of the American Educational Research Association, Chicago.

Fullan, M. G. (1991). *The new meaning of educational change*. New York: Teachers College Press.

Hanushek, E. A. (1994). *Making schools work: Improving performance and controlling costs*. Washington, D.C.: Brookings Institution.

Hargreaves, D. H., & Hopkins, D. (1994). *Development planning for school improvement*. New York: Cassell.

Hutchinson, B. (1993). The effective reflective school: Visions and pipedreams in development planning. *Educational Management and Administration, 21*(1), 4-18.

Levine, D. U., & Leibert, R. E. (1987). Improving school improvement plans. *The Elementary School Journal, 87*(4), 397-412.

Lezotte, L. W., & Bancroft, B. A. (1985). School improvement based on effective schools research: A promising approach for economically disadvantaged and minority students. *Journal of Negro Education, 54*(3), 301-312.

Lindblom, C. E. (1959). The science of muddling through. *Public Administration Review, 19*(Spring), 79-88.

Little, J. W. (1982). Norms of collegiality and experimentation: Workplace conditions of school success. *American Educational Research Journal, 19*(3), 325-340.

Malen, B., & Muncey, D. (1999). Creating "a new set of givens?": The impact of state activism on school autonomy. *Draft copy of chapter for the AEFA Yearbook* .

Maryland State Department of Education Division of Planning Results and Information Management. (1997). *Criteria for Reconstitution*: Maryland State Department of Education.

McDonald, J. P. (1996). *Redesigning school: Lessons for the 21st century*. San Francisco: Jossey- Bass.

McLaughlin, M.W. (1990). The RAND change agent study revisited: macro perspectives and micro realities. *Educational Researcher, 19*, 11-16.

McLaughlin, M. W., & Talbert, J. E. (1993). Contexts that matter for teaching and learning: Strategic opportunities for meeting the nation's educational goals: Center for Research on the Context of Secondary School Teaching.

Meier, D. (1995). *The power of their ideas: Lessons for America from a small school in Harlem*. Boston: Beacon Press.

Meyer, J., & Rowan, B. (1978). The structure of educational organizations. In M. W. Meyer (Ed.), *Environments and organizations* (pp. 78-109). San Francisco: Jossey-Bass.

Mitchell, S. M. (1995). Performance management in education. In S. B. Bacharach & B. Mundell (Eds.), *Images of schools structures and roles in organizational behavior* (pp. 201-238). Thousand Oaks: Corwin Press.

Morrison, K. (1998). *Management theories for educational change*. Thousand Oaks: Sage.

Muncey, D. E., & McQuillan, P. J. (1996). *Reform and resistance in schools and classrooms: An ethnographic view of the coalition of essential schools*. New Haven: Yale University Press.

Neale, D. C., Bailey, W. J., & Ross, B. E. (1981). *Strategies for school improvement: Cooperative planning and organization development*. Boston: Allyn and Bacon.

Newmann, F. M., King, M. B., & Rigdon, M. (1997). Accountability and school performance: Implications from restructuring schools. *Harvard Educational Review, 67*(1), 41-74.

Nias, J., Southworth, G., & Campbell, P. (1992). *Whole school curriculum development in primary school*. Washington, D.C.: The Falmer Press.

O'Day, J. A., & Smith, M. S. (1993). Systemic reform and educational opportunity. In S. H. Fuhrman (Ed.), *Designing coherent education policy: Improving the system* (pp. 250-312). San Francisco: Jossey-Bass.

Olsen, J. (1976). Reorganization as a garbage can. In March, J. G. & J. Olsen (Eds.), *Ambiguity and choice in organizations* (pp. 314-336). Bergen, Norway: Universitetsforlaget.

Patterson, J. L., Purkey, S. C., & Parker, J. V. (1986). *Productive school systems for a nonrational world*. Alexandria: Association for Supervision and Curriculum Development.

Rowan, B. (1986). *Rationality and reality in instructional management: Results from a survey of districts in the Far West*. San Francisco: Far West Laboratory for Educational Research and Development.

Sarason, S. (1990). *The predictable failure of educational reform*. San Francisco: Jossey-Bass.

Schein, E. (1991). *Organizational culture and leadership*. San Francisco: Jossey-Bass.

Smith, M. S., & O'Day, J. (1991). Systemic school reform. In S. H. Fuhrman & B. Malen (Eds.), *The politics of curriculum and testing* (pp. 233-267). London: The Falmer Press.

Stone, M. M., & Brush, C. G. (1996). Planning in ambiguous contexts: The dilemma of meeting needs for commitment and demands for legitimacy. *Strategic Management Journal, 17*, 633-652.

Tucker, M. S., & Codding, J. B. (1998). *Standards for our schools: How to set them, measure them, and reach them*. San Francisco: Jossey-Bass.

Wallace, M. (1994). Towards a contingency approach to development planning in schools. In D. H. Hargreaves & D. Hopkins (Eds.), *Development planning for school improvement* (pp. 150-161). New York: Cassell.

Weick, K. (1978). Educational organizations as loosely coupled systems. *Administrative Science Quarterly, 23*(December), 541-552.

Weiner, B. (1986). *An attributional theory of motivation and emotion*. New York: Springer-Verlag.

Wise, A. (1977). Why educational policies often fail: The hyperrationalization hypothesis. *Journal of Curriculum Studies, 9*(1), 43-57.